Understanding the Impact of Dementia on Brain Structure

1. Introduction

Our research focuses on understanding how dementia impacts two critical brain volume measures: normalized whole brain volume (nWBV) and estimated total intracranial volume (eTIV). We aim to explore how these volumes differ across individuals classified into various dementia groups and how these measures change over successive MRI scans. Specifically, the study investigates whether there are significant variations in nWBV and eTIV among dementia groups and across different visits, along with the potential interaction effects of dementia classification and the timing of MRI scans on these brain volumes.

Research Question 1: Does the normalized whole brain volume (nWBV) vary significantly among different dementia groups (determined by the CDR) and across multiple visits (Visit), and is there an interaction effect between group classification and visit time on nWBV?
Research Question 2: Does the estimated total intracranial volume (eTIV) vary significantly among different dementia groups (determined by CDR) and across multiple visits (Visit), and is there an interaction effect between group classification and visit time on eTIV?

2. Data Cleaning and Data Wrangling

The dataset, "INF2178_A4_data.csv," which consists of 294 entries and 16 columns, captures a wide range of variables essential for analyzing the progression and effects of dementia through MRI scans. Initial examination indicates that the dataset is nearly ready for in-depth analysis, needing only slight adjustments to align with our research goals.
CDR: Clinical Dementia Rating
Visit Time: Tracks when MRI scans are done, crucial for seeing how brain volume changes over time.
nWBV: As a main focus of our study, this measure lets us examine brain volume variations across dementia groups and over time.
eTIV: Also a core part of our analysis, it looks at changes in total brain volume by dementia group and visit.
Subject ID: A vital identifier for tracking each participant's brain volume data across all scans.

3. Exploratory Data Analysis (EDA)

We proceeded with a comprehensive EDA to leverage insight that could potentially lead to interesting research questions.We started by describing our quantitative data in Figure 1 below.

	eTIV	nWBV	CDR	Age
count	291	291	291	291
mean	1478.244	0.732	0.284	76.323
std	176.885	0.037	0.342	7.538
min	1106	0.646	0	60
25%	1345.5	0.703	0	71
50%	1461	0.732	0	76
75%	1569	0.756	0.5	81
max	2004	0.837	1	96

Figure 1: *Dataset quantitative Data statistics*

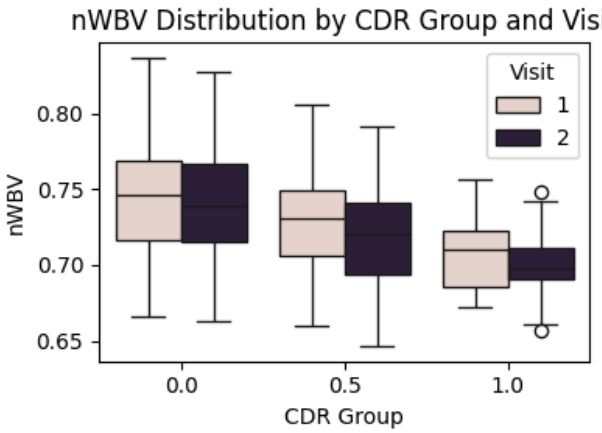

Figure 2: *Distribution of nWBV*

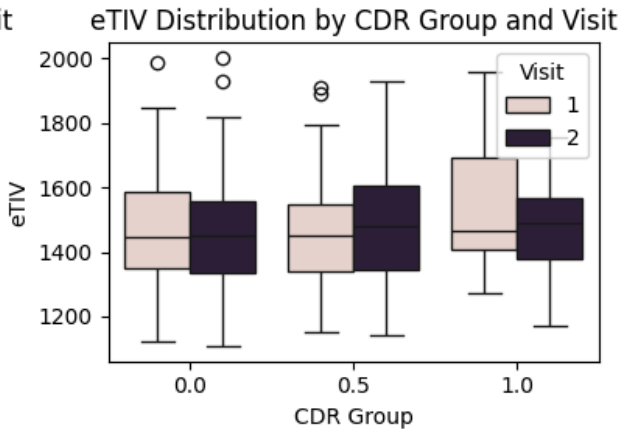

Figure 3: *Distribution of eTIV*

The box plots illustrate that the distribution of nWBV decreases as the CDR increases, which is expected since a higher CDR score indicates more severe dementia, often associated with greater brain atrophy.

The distribution of eTIV does not show a consistent pattern across CDR scores, suggesting that while brain volume decreases with dementia severity, the overall size of the intracranial cavity (which can be affected by many factors) does not change in the same way.

There's variability across visits within the same CDR groups for both nWBV and eTIV, hinting at the progressive nature of dementia but also the influence of other factors over time.

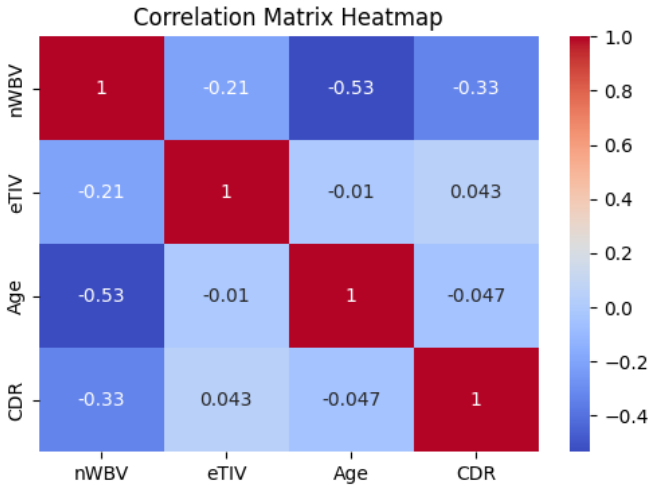

Figure 4: *correlation matrix*

1. nWBV and Age: There is a significant negative correlation (-0.54) between nWBV and Age. This indicates that as age increases, the normalized whole brain volume tends to decrease, which is consistent with the general understanding that brain atrophy occurs with aging.

2. nWBV and CDR: A negative correlation (-0.33) exists between nWBV and CDR, suggesting that higher CDR scores, which indicate greater dementia severity, are associated with lower brain volumes.

3. **CDR and Age**: CDR has a very weak positive correlation (0.011) with Age, which is somewhat surprising because one might expect that dementia severity could increase with age. This could suggest that within this dataset, dementia severity as captured by CDR scores is not strongly dependent on age, or it could indicate that the sample is not representative of the general trend.

4. Mixed-Effect ANOVA Result for Research question 1

Figure 5: *Mix-effect ANOVA table for nWBV*

Source	SS	DF1	DF2	MS	F	p-unc	np2	eps
Dementia_Group	0.040	2	115	0.020	7.723	7.126207e-04	0.118	NaN
Visit	0.004	1	115	0.005	73.063	6.137789e-14	0.389	1.0
Interaction	0.000389	2	115	0.000194	2.928	5.748719e-02	0.048	NaN

Dementia_Group: The p-value for the Dementia_Group factor is p=7.126207e−04, which is less than the conventional alpha level of 0.05, indicating that there are statistically significant differences in the normalized whole brain volume (nWBV) across the different dementia groups.

Visit: The p-value for the Visit factor is p<.001 (precisely 6.137789e−14), which is much less than the alpha level of 0.05. This shows a statistically significant difference in nWBV between the different visits. The large F-value (F = 73.063) suggests a strong effect

Interaction: The interaction effect has a p-value of 5.748719(precisely 0.048), which is very close to the alpha level of 0.05, suggesting a marginal interaction effect between the dementia group and the visit on nWBV.

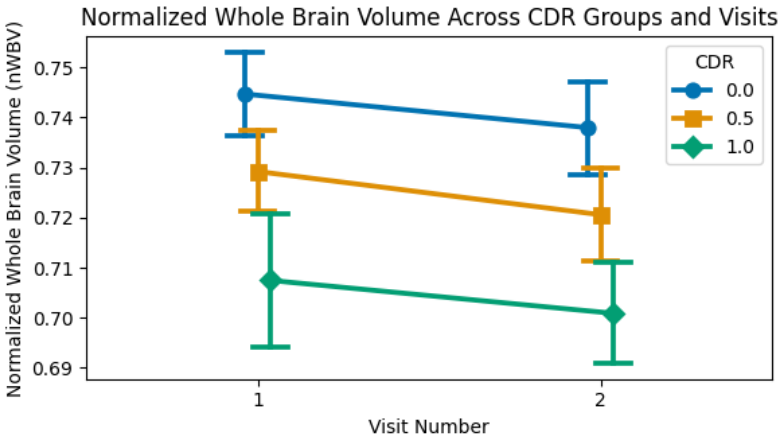

***Figure 6:** Output plots for nWBV*

Difference in nWBV across CDR Groups: Each CDR group (0.0, 0.5, and 1.0) has a distinct range of nWBV values. It appears that the nWBV is highest for the CDR 0.0 group and lowest for the CDR 1.0 group. This corresponds with what you might expect since higher CDR values indicate more severe dementia, which is often associated with greater brain atrophy.

Changes in nWBV over Visits: For each group, there seems to be a decline in nWBV from Visit 1 to Visit 2. This is consistent with the progression of atrophy over time, which may be particularly relevant in the context of dementia.

Test For Assumption:

- **Normality Test:**
  For all dementia groups (CDR 0.0, 0.5, 1.0) and both visits, the brain volume data is normally distributed (since p > 0.05).
- **Homogeneity of Variances Test (Levene's Test):**
  The variances of brain volume p = 0.0135 are not the same across dementia groups (since p < 0.05). This could affect the ANOVA results.
- **Mauchly's test of sphericity**: This test is applicable only if within-subject factor has more than two levels

5. Mixed-Effect ANOVA Result for Research question 2

***Figure 7:** Mix-effect ANOVA table for eTIV*

Source	SS	DF1	DF2	MS	F	p-unc	np2	eps
Dementia_Group	24593.95	2	115	12296.975	0.19	0.827	0.00329	NaN
Visit	5034.322	1	115	5034.322	8.553	0.00416	0.0692	1.0
Interaction	15135.993	2	115	7567.997	12.858	0.000009	0.183	NaN

Dementia_Group:SS (Sum of Squares) is quite large, indicating variability due to the dementia group factor.DF1 and DF2 are the degrees of freedom associated with the dementia group factor and within-subjects error, respectively. F is the F-statistic value, which is relatively small and corresponds to the ratio of variance explained by the group to the variance within the groups. p-value (p-unc) is not significant (p > 0.05), suggesting that there is no significant effect of the dementia group factor on eTIV.

Visit:SS is smaller compared to the dementia group, indicating less variability due to the visit factor. DF1 and DF2 show degrees of freedom for the visit factor and within-subjects error. F is higher than for Dementia_Group, suggesting a greater effect size. p-value is significant (p < 0.05), indicating a statistically significant effect of visit number on eTIV.

Interaction (Dementia_Group * Visit): SS is substantial, suggesting there is considerable variability due to the interaction between dementia group and visit number. DF1 and DF2 are the degrees of freedom for the interaction and error, respectively. F is very high, indicating a strong interaction effect. p-value is extremely significant ($p < 0.0001$), suggesting a very significant interaction effect between the dementia group and visit number on eTIV.

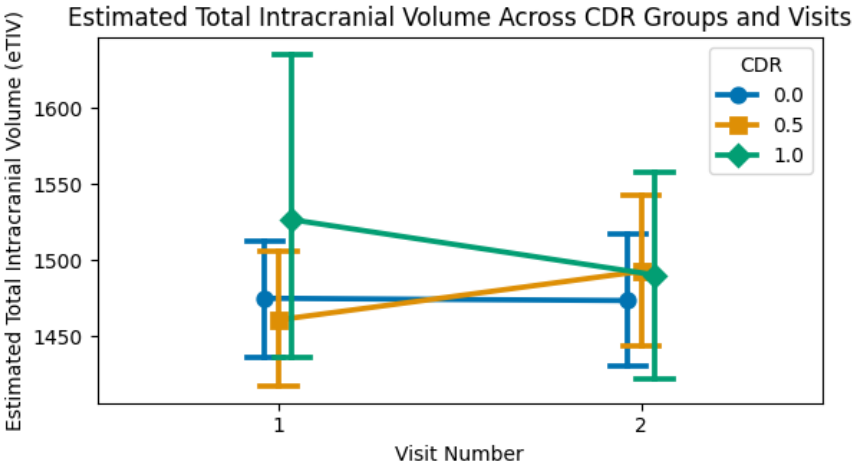

***Figure 8:** Output plots for eTIV*

Trends: The eTIV appears to decrease from the first to the second visit across all CDR groups. This decrease might suggest atrophy or other changes associated with the progression of dementia, but this is speculative without further clinical context.

Error Bars: The plot includes error bars that represent the variability or confidence interval around the mean eTIV for each group at each visit. The length of these bars indicates that there is considerable variability within each group at each time point, especially in the CDR 0.0 group during the first visit.

Inter-Group Differences: At the first visit, the CDR 0.0 group appears to have a slightly higher mean eTIV compared to the other groups, though the large error bars indicate overlap. By the second visit, the mean eTIV values of all groups are more similar, suggesting that the difference in brain volume between the groups may be less distinct over time or that the groups converge.

Test For Assumption:

- **Normality Test:** For the CDR 1.0 group, the data appears to be normally distributed ($p = 0.252424 > 0.05$), while for the CDR 0.0 and CDR 0.5 groups, the normality assumption does not hold ($p < 0.05$). Similarly, for both visits 1 and 2, the normality assumption is violated ($p < 0.05$), which suggests that the data for these groups and visits are not normally distributed.
- **Homogeneity of Variances Test (Levene's Test):**Levene's test for equality of variances shows that variances are equal across the different dementia groups ($p = 0.733472 > 0.05$). This suggests that the assumption of homogeneity of variances is met for conducting an ANOVA.
- **Mauchly's Test of Sphericity:**Mauchly's test of sphericity is not applicable here since the within-subject factor, 'Visit', only has two levels.

These tests of assumptions indicate that while the data may not be normally distributed for all dementia groups and visits, the variances are homogeneous across groups.

6. Statistical power for t-tests and appropriate sample size

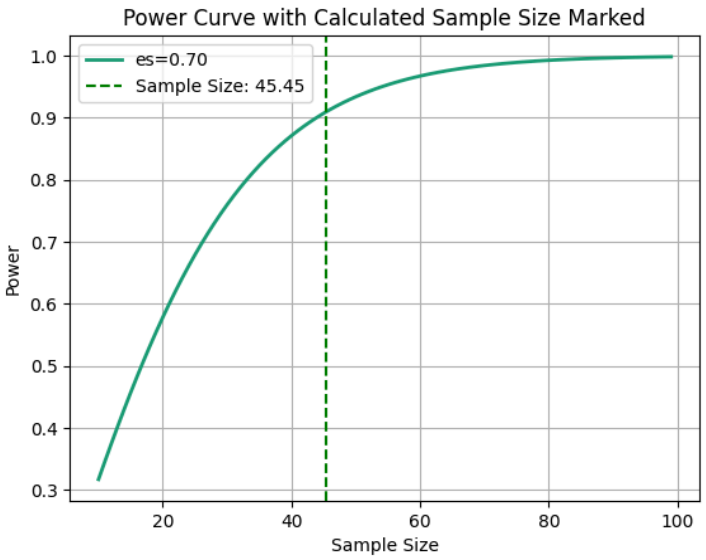

***Figure 9:** Power plot for t test*

The appropriate sample size needed in each group is 45.451, approximately 46 participants.

7. Conclusion

The effects and interactions observed in nWBV and eTIV indicate that both measurements are influenced by the severity of dementia and change over time. The interaction effects suggest that these changes are not uniform across different levels of dementia severity, highlighting the complexity of how dementia progresses over time. The failure to meet certain ANOVA assumptions, especially in the eTIV analysis, suggests that the results should be interpreted with caution and may benefit from further analysis using alternative statistical methods that are less sensitive to these assumptions.